Abstract of the Disclosure

In a non-volatile memory, a programming cycle consists of the following phases: high voltage charging up, programming pulse, and discharge. The actual programming process only takes place in the programming pulse phase. Several break points are defined relative to elapsed time and introduced in the programming pulse phase. Upon receiving a suspend request, the programming operation will advance to the next break point, then discharge the high programming voltage and go to a suspend state. A separate counter is used to monitor the break points so that elapsed non-programming time can be deducted from the total programming pulse time when the programming operation is resumed. By doing so, the device can handle frequent suspend and resume requests. Since the total time duration in the programming pulse phase is equal for the programming operation with and without suspend and resume requests, the programming proceeds efficiently to completion.